

# Data Mining in Evaluating the Impact of Perceived Trust in the Consumption of Safe Foods in Vietnamese Households: The Case of Vegetables in Hanoi

**Tran Thi Thu Ha**

*National Economics University, Vietnam*

*Email: hattthhkt@neu.edu.vn*

**Nguyen Thi Minh**

*National Economics University, Vietnam*

*Email: minhnt@neu.edu.vn*

**Le Thi Anh**

*National Economics University, Vietnam*

*Email: leanhtoankt@neu.edu.vn*

**Kieu Nguyet Kim**

*Hanoi University of Industry, Vietnam*

*Email: kieu.kim@hau.edu.vn*

Received: 27 July 2017 | Revised: 12 January 2018 | Accepted: 27 February 2018

## **Abstract**

*Food safety is as much of a concern to Vietnamese citizens as it is to the public authorities. As safe vegetables are classified as credence goods, the markets of which exhibit a high level of information asymmetry between the buyers and the suppliers. As such, making the market for safe vegetables become more transparent and grow sustainably is a must, but not an easy task. In this paper, we use a Kernel regression method to discover the main determinants of consumers' decisions for the consumption of "safe" vegetables with more focus on perceived levels of trust. The result shows that apart from other traditional factors, perceived trust is an important determinant of consumers' decisions. However, the data shows that consumers put more trust in un-verified factors such as "store's reputation" or "label" and much less on formal factors such as "government certificates". This result raises some alarm as other studies show that without trusted involvement from the Government, signals from suppliers, such as labeling are not reliable.*

**Keywords:** Kernel regression; perceived trust; safe vegetables.

**JEL code:** C14, D12.

---

## 1. Introduction

Vegetables are considered to be a very important ingredient in the daily diet, especially for people who live in an agricultural country like Vietnam (Chen, 2007). With an alarming situation of vegetable safety, the demand for safe vegetables is increasing. The supply system for safe vegetables has been developed quite strongly. In 2008, the Government, together with the Ministry of Agriculture and Rural Development (MARD), developed and implemented the VietGAP program, which aims at providing assistance for farmers who grow safe vegetables. Along with the supermarkets, there are many stores that sell safe vegetables in big cities. Selling safe vegetables occurs in many places in the big cities along with the supermarket system in order to meet the increasing demand from residents.

However, we observe a paradox in the market for safe vegetables. The gap between the demand side and the supply side for safe vegetables is consistently large. On the one hand, growers of safe vegetables find it difficult to sell their products to people in need<sup>1</sup>. In many cases they have to sell their products to wholesalers as if the products were conventional vegetables of a low price. On another hand, people who live in urban areas are struggling to find vegetables sellers who they can trust about the safety of their product. As a result, many people in big cities have to protect themselves by growing vegetables themselves on the rooftops or balconies of their houses at a very high cost and with a high time consumption. From the supply side, the programs promoting safe vegetable planting supported by the Government such as the “Safe vegetables program” in 1995

(Mergenthaler et al., 2009), or more recently, the VietGap program<sup>2</sup> implemented since 2008, have not gained much trust from customers. After 10 years of establishment, VietGap covers only 0.4% of the total area for growing vegetables<sup>3</sup>. Farmers are reluctant to plant safe vegetables and customers are reluctant to buy products marked as “safe vegetables”. According to Alexander (2014), in 2014, safe vegetables accounted for only 3.2% of the total expenditure for vegetables of Hanoi people.

One of the main reasons for the paradox is the information asymmetry in the market for safe vegetables. While sellers may know about the safety of the vegetables, buyers do not, even after consuming them. In other words, safe vegetables can be classified as credence goods: goods for which expenditure is based mainly on consumers’ perceived trust about their quality (McCluskey, 2000). The theory of information asymmetry is proposed by Akerlof (1970) (a Nobel prize winner in economics in 2002). The theory states that information asymmetry will render the market to move away from its optimal status; and severe asymmetry may even lead to a market collapse. High quality products are often produced with a higher cost, but if customers can not distinguish them from low quality ones of a lower cost then there is no motivation for producing high quality products, and gradually there are no longer high quality products in the market. In order to solve the problem, Spencer (1973) proposed the signaling theory; and Stiglitz (1975) proposed the screening theory. While the latter approaches the problem from the demand side, encouraging users to screen for more information about products, the former pays attention to the sup-

---

ply side, which asks sellers to provide more information to potential customers.

Studies about behavior of consumers in the food market often focus on consumer demand, willingness to pay, or determinants of willingness to pay (Chih-Ching Teng and Yu-Mei Wang, 2015; Gracia and Magistris, 2008, Janssen and Hamm, 2012). When it comes to *credence goods* such as organic foods or safe foods, studies are interested in the role of signaling factors, including labels, certificates, price, or consumers' trust. In other words, besides the traditional factors, consumers' perceived trust towards signals is of great interest in many studies in the field. One of the lines is the study of Chih-Ching Teng and Yu-Mei Wang, (2015) about the demand of Taiwan people for organic foods. The authors found that consumer trust is the most important determinant when making decision buying or not buying an organic food. The same conclusion is also found in the study of Xu and Lu (2010) which examines the rank of determinants of Chinese consumers' decisions for safe foods, with pork as a case study. In this study, the authors used a logit model with random coefficients on a sample size of 420. The result shows that a government certificate is the factor that Chinese people trust most, follows by other certificates, information about the production field and producers, and the last is labels with other information.

In industrialized countries, where state surveillance as well as inspection systems are well functioning, customers still require guarantees from the government in order to trust the signals provided by suppliers. For example, the study of Roosen and Lusk (2003) of beef demand in Britain, USA and Australia shows that

people in these countries very much desire that labeling is mandatory by the government, even though this may lead to a 2% increase in beef price. These results are consistent with many other findings, including that by McCluskey (2000) when studying asymmetric information in the market for organic foods. McCluskey (2000) concludes that with credence goods, without quality control measures from government, signals provided by suppliers may be invalid. Moreover, Roosen et al. (2003) showed that consumers put more trust in the signals provided by mass production suppliers than by retailers

To sum up, studies of the market for safe foods agree on the important role of perceived trust of signals provided by both government and suppliers. Also, signals provided by wholesalers gain more trust than signals provided by small sellers. In a developing country like Vietnam, where the public inspection system has not been well functioning, and the distribution system is still rather primitive, where foods and vegetables are distributed mostly by individual sellers in street markets, how to control the safety of vegetables as well as to build up consumers trust is not an easy task.

In Vietnam, there have been a few studies about demand for vegetables, such as the study by Nguyen Thi Hong Trang (2016). However, these studies either focus on the procedure for growing safe vegetables (supply side), or basic statistical analysis of the status of the market, and have not paid attention to consumers' behaviors (demand side). Other studies on asymmetric information such as Nguyen Thi Minh and Hoang Bich Phuong (2012), Nguyen Thi Minh et al. (2014). However, these studies are

concerned with the health insurance market and the stock market. Hence, we hope that this work will contribute to the literature on customer behavior in the market of safe vegetables in Vietnam. The structure of the work is as follows: the next section introduces the Kernel regression method, Section 3 presents data and empirical results, Section 4 concludes and proposes some policy recommendations.

## 2. Non-parametric Kernel regression

For the sake of the presentation, assume that the research interest is the relationship between a dependent variable  $Y$  and an explanatory variable  $X$ :

$$E(Y|X) = m(X) \quad (1.1)$$

In which  $m(X)$  is some function of  $X$ . With a parametric approach,  $m(\cdot)$  is assumed to take some specific form, for example,  $m(\cdot)$  could be a linear function:

$$E(Y|X) = \beta_1 + \beta_2 X \quad (1.2)$$

Then parametric methods such as OLS, ML or GMM can be applied for parameter estimation. The estimates of  $\beta_1, \beta_2$  from a parametric approach are often easy to interpret. However, if  $m(\cdot)$  is misspecified then the estimators are biased and inconsistent, leading to a misleading conclusion and incorrect inference. In many cases, imposing a specific function form for  $m(\cdot)$  could be hard, then a non-parametric approach is a good alternative. The paper will apply Kernel regression to estimate (1.1). This is a modern approach based on Kernel function, as follows.

We have:

$$\begin{aligned} m(x) &= E(Y|X=x) = E(Y|_{X=x}) \\ &= \int_R y f(y|x) dy \quad (1.3) \end{aligned}$$

Where  $f(y|x)$  is the density function of  $Y$  conditional on  $X$ . The non-parametric method that uses the Kernel density function to estimate (1.3) is named as the Kernel regression method.

Some popular Kernel functions in regression include: The Epanechnikov function

$$K(z) = \frac{3}{4}(1-z^2)1(|z| \leq 1)$$

with  $1(|z| \leq 1)$  is the index function, or normal Kernel:

$\ddot{u}(\cdot) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$  for continuous variables, and Aitchison or Aitken for nominal variables.

Two common methods used in Kernel regression: local constant method and local linear method. The former is proposed by Nadaraya (1964) and Watson (1964) and are known as N-W (Nadaraya-Watson):

$$\hat{m}_h(x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)} \quad (1.4)$$

In which  $K_h(\cdot)$  is Kernel density function with bandwidth  $h$ . Under regular conditions of Kernel function, Nadaraya (1964) proved that (1.4) is a consistent estimator of  $m(x)$ . This estimator, however is often biased at the boundary and where the distribution is not so homogenous.

The local – linear method proposed by Li and Racine (2004) overcomes the bias problem in the N-W method. The idea of the method can be briefly outlined as follows: within a neighborhood of  $X_0$ , it assumes that  $Y$  is a linear function of  $X$  within some neighborhood of  $X_0$  instead of assuming constant  $Y$  as in N-W.

More specifically, at each point  $x$ , we find coefficient vectors  $\alpha(x)$ ,  $\beta(x)$  such that:

$$\text{Min}_{\alpha(x),\beta(x)} \sum_{i=1 \rightarrow N(x)} \frac{(y_i - \alpha(x) - (x_i - x)\beta(x))^2}{N(x)} K_h(x_i - x) \quad (1.5)$$

In which the summation is taken over the observation  $x_i$ :  $|x_i - x| \leq h$  with chosen bandwidth  $h$ . In this paper, we use the local – linear method.

### 3. Model and empirical results

This section will present the results from Kernel regression estimation using a primary data set. For a robustness check, we compare the results with the estimates received by parameter estimation.

#### 3.1. Data

The dataset used in this paper was collected by the authors. The data collection was con-

ducted as follows: the sample was selected according to a convention rule so that it covered different components of housing characteristics (apartments and other residential areas) and workplaces (public units, schools, private sectors). The investigator went from door to door to distribute questionnaires and came back one week later to collect them. Questionnaires were constructed based on a literature review and pilot survey which consisted of 50 people randomly chosen. The 700 questionnaires were distributed of which 54 had missed answers leaving 646 valid responses for usage in the calculation. Basic statistics of the sample are in Table 1.

Perceived trust: how much consumers trust the seller — taking values from 1 (very trusting) to 5 (less trusting). We expected that the

**Table 1: Sample statistics**

| Variables       | Values  | No of observations | %     |
|-----------------|---------|--------------------|-------|
| Perceived trust | 1       | 71                 | 5.5   |
|                 | 2       | 573                | 44.35 |
|                 | 3       | 500                | 38.7  |
|                 | 4       | 118                | 9.13  |
|                 | 5       | 30                 | 2.32  |
| Education       | 1       | 394                | 30.5  |
|                 | 2       | 792                | 61.3  |
|                 | 3       | 106                | 8.2   |
| Google          | 1       | 564                | 43.65 |
|                 | 2       | 430                | 33.28 |
|                 | 3       | 298                | 23.07 |
| Gender          | Female  | 1098               | 85    |
|                 | Male    | 194                | 15    |
| Age             | <25     | 108                | 8.04  |
|                 | 25<÷<40 | 614                | 47.5  |
|                 | >40     | 570                | 44.1  |
| Children        | 1       | 664                | 51.39 |
|                 |         | 628                | 48.61 |

Source: Calculated from the surveyed data.

---

more a consumer trusts a seller, the more he/she purchases products from that seller.

*Education:* the highest degree of education, a categorical variable, taking a value 1 for lower than bachelor degree, 2 for having a bachelor degree, and 3 for post graduate. This variable indicates the attitude towards the risk of having unsafe vegetables. Our hypothesis is that higher educated people care more about the safety of their diet.

*Google:* how often the respondents search for information about safe vegetables: a categorical variable, taking a value of 1 for rarely, 2 for often, and 3 for very often. This variable represents the extent a person cares about safety.

*Gender:* 1 for female, 0 for male. We expect that female people may be more risk averse than their male counterparts.

*Children:* 1 for having children under 6 years of age, 0 for otherwise. Families with young children often pay more for safe foods.

Some statistics in the sample may not represent the structure of the population of Hanoi. In the sample, 84,9% respondents are female, which is too large a proportion compared with the actual percentage of females in Hanoi. However, in Vietnam, people who take care of food and vegetables for their family are mainly female, so this differential is appropriate.

### **3.2. Model and non-parametric estimation results**

Our model takes the form of:

$$buy = m(trust, consumption, ageq, educ, concern, type) \quad (2.1)$$

In which:

*Buy:* the percentage of budget used for safe

vegetables in the total budget for vegetables, the dependent variable.

*Trust:* the consumer's perceived trust towards the shop that the vegetables are safe. The higher the trust is, the more likely the consumer will buy at the shop; this is the main variable in our analysis

*Consumption:* adjusted expenditure for vegetables per head, which is per head expenditure on vegetables. As the price of safe vegetables is higher than for normal vegetables, we need to adjust for this in order to estimate the demand for vegetables. We argue that vegetables can be classified as necessary goods for Vietnamese people, hence the demand for vegetables is assumed to be met - the point is the choice between the normal vegetables with a lower price and the safe ones with a higher price. The demand for vegetables may be heterogeneous among households, hence besides income, the consumption may reflect the household purchasing capacity.

*Educ:* a dummy variable, taking a value of 1 for people with high school or less, 2 for bachelor degree holders, and 3 for post graduates. This variable reflects the attitude towards risk as well as recognition of the capacity of households for risk.

*Google:* a dummy variable, taking a value of 1 for people who search for information about food safety very rarely, 2 for often, and 3 for very often. This variable is included to indicate how much the household cares about food safety.

*Ageq* and *gender* are age group and gender and are demographic characteristics that may affect behavior in consuming vegetables > The elderly or females may care more about health



**Table 2: Basis statistics of variables**

| Variable | buy  | trust | Consumption | educ | google | type | children |
|----------|------|-------|-------------|------|--------|------|----------|
| mean     | 0.23 | 2.58  | 48.12       | 1.78 | 1.79   | 1.50 | 0.49     |
| min      | 0    | 1     | 4.571429    | 1    | 1      | 1    | 0        |
| max      | 1    | 5     | 220         | 3    | 3      | 2    | 1        |
| sd       | 0.29 | 0.82  | 28.48       | 0.58 | 0.79   | 0.50 | 0.50     |
| N        | 1292 | 1292  | 1292        | 1292 | 1292   | 1292 | 1292     |

Source: Calculated from surveyed data.

then the others.

*Type*: a dummy variable, taking a value of 1 for supermarkets, and 0 for other shops that sell safe vegetables. Although prices are very much the same between the two types, the attractiveness may differ. Shops may have a more intimate relationship with their customers.

As mentioned, prices are much the same between the two types of sale outlets, hence are not included in the model.

The estimation of (2.1) using non-parametric Kernel regression is conducted through 3 steps:

Step 1: Testing of the parameter vs. non-parameter function form. The test used is proposed by Hsiao et al. (2007). The test result based on bootstrapping over 399 times (in Appendix 1) yields a probability  $p = 0.07$ , implying that a non-parametric model is more appropriate. The next step will be the estimation of the non-parametric model.

Step 2: For the result for non-parametric Kernel regression to be reliable, we need to determine the bandwidth for each variable in the model. This is based on the cross-validation method, which is to find a bandwidth  $h$  that minimizes forecast error:

$$CV(h) = \frac{1}{n} \sum_{i=1 \rightarrow n} \{Y_i - \hat{m}_{h(-i)}(x_i)\}^2$$

Where  $\hat{m}_{h(-i)}(x_i)$  is  $\hat{m}_h(x_i)$  calculated after removing  $x_i$  and standardized so that the total weight equals to 1 (Alexander, 2014, p.70).

The chosen bandwidth will be used next to estimate, using Kernel regression.

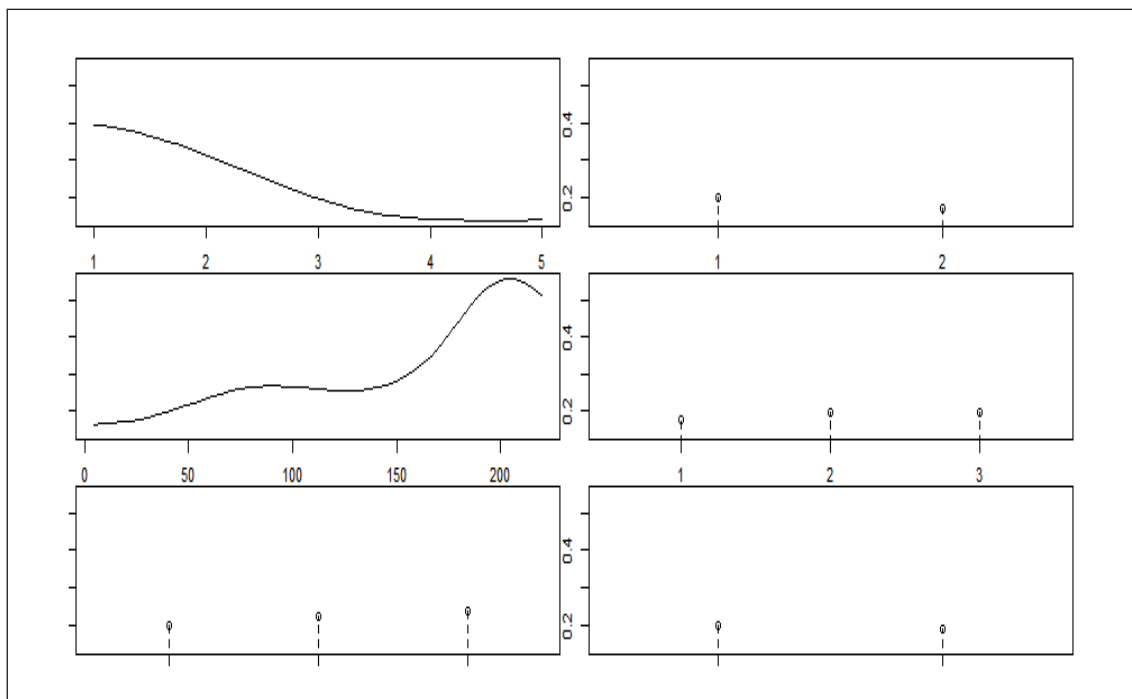
Step 3: Testing about the statistical significance of coefficients using the bootstrap method. Test result shows that (Appendix 2) all variables are statistically significant at 1% and 5% apart from age and gender. The marginal effects are reported in Figure 1.

Figure 1 depicts the marginal impact of: trust, type, consumption, educ, google, children on the share of spending on safe vegetables (respectively in the order from left to right, from top to the bottom).

It can be seen from Figure 1 that the result is consistent with the expectation, in which *trust* is negatively related with proportion with safe food consumed (recall that  $trust = 1$  is for very trustworthy, 5 for not at all trustworthy). People tend to buy more at supermarkets instead of special shops. Consumption, representing household purchasing capacity, is positively related to the proportion of safe food consumed. More specifically:

- The impact of *trust* is very clear, at a high level of trust ( $trust = 1$ ), the proportion of safe

**Figure 1: Marginal effect of variables on percentage of spending on safe vegetables**



Source: Calculated by authors using surveyed data in R software.

vegetables consumed to total vegetables is about 0.4, at *trust* = 2, the number is still large at 0.3. At a low level, *trust* = 4 or *trust* = 5, the number is very low. Furthermore, the impact is not in a linear form, which is to reaffirm that a non-parametric method is more suitable than a parametric one.

- Regarding variable *type*: The proportion of safe vegetables bought at supermarkets is larger than that at specialist shops. This result is consistent with the fact that people may tend to go shopping more at supermarkets for more convenience where they can buy many things at the one place.

- Regarding *education*, the difference in the proportion of safe vegetables among education groups is also statistically significant. How-

ever, the difference is not large, implying that people worry about food safety regardless of their level of knowledge.

- The variable *Google* also has a clear impact: the more people are concerned about safety, the more they pay for safe vegetables.

- Having children or not does not impact on the proportion of safe vegetables consumed; this result may be consistent with the above statistical analysis: people are quite concerned about food safety.

### 3.3. Robustness check

To do the robustness check, we compare the model above with a parametric model.

We consider the following parametric model:



**Table 3: Estimated result for the parametric model**

| Buy             | Coef. | Std. Err. | T     | P>t  | [95% Conf. | Interval] |
|-----------------|-------|-----------|-------|------|------------|-----------|
| Trust           |       |           |       |      |            |           |
| 2               | -0.12 | 0.04      | -3.40 | 0.00 | -0.19      | -0.05     |
| 3               | -0.19 | 0.04      | -5.24 | 0.00 | -0.26      | -0.12     |
| 4               | -0.25 | 0.04      | -5.99 | 0.00 | -0.34      | -0.17     |
| 5               | -0.24 | 0.06      | -3.88 | 0.00 | -0.36      | -0.12     |
| Consumption     | 0.00  | 0.00      | 3.74  | 0.00 | 0.00       | 0.00      |
| Gender          | -0.01 | 0.02      | -0.43 | 0.67 | -0.05      | 0.03      |
| Educ            |       |           |       |      |            |           |
| Bachelor degree | 0.08  | 0.02      | 4.36  | 0.00 | 0.04       | 0.11      |
| Post graduate   | 0.11  | 0.03      | 3.34  | 0.00 | 0.04       | 0.17      |
| Google          |       |           |       |      |            |           |
| 2               | 0.04  | 0.02      | 2.10  | 0.04 | 0.00       | 0.07      |
| 3               | 0.08  | 0.02      | 3.75  | 0.00 | 0.04       | 0.12      |
| Type            |       |           |       |      |            |           |
| Stores          | -0.09 | 0.02      | -6.05 | 0.00 | -0.13      | -0.06     |
| cons            | 0.32  | 0.06      | 5.39  | 0.00 | 0.20       | 0.43      |

$$buy = \beta_0 + \beta_1 trust + \beta_2 consumption + \beta_3 type + \beta_4 children + \beta_5 google + \beta_6 educ + u$$

The estimated result is reported in Table 3.

To compare the two models, we process as follows:

We divide the data set into 2 subsets, the first one consists of 1000 observations, and the second 292 observations used for model evaluation. We run both models using the first set, and evaluate the models in both the evaluation set and the whole set. The comparison is based on R<sup>2</sup> and Mean square error (MSE), as in Table 4.

Table 4 shows that the result from the non-parametric model is better.

#### 4. Conclusion and recommendation

From the analysis, it can be seen that perceived trust is critical in consumers' decisions for purchasing safe vegetables. When trust is from neutral downward, people spend very little on safe vegetables (after controlling for other factors). This implies that enhancing trust is a key to the expansion of demand for safe vegetables.

Furthermore, the data show that consumers place most trust on labels and the store's reputation (Minh et al., 2017), both of which are difficult for them to verify. At the same time, a "government certificate" which is a formal

**Table 4: Comparison of the parametric model and non-parametric model**

|                | R2         |         | MSE        |          |
|----------------|------------|---------|------------|----------|
|                | Parametric | Non-par | Parametric | Non- par |
| Whole set      | 0.1057     | 0.24984 |            |          |
| First set      | 0.08535    | 0.23652 |            |          |
| Evaluation set |            |         | 0.0839     | 0.07251  |

---

factor, receives a low level of trust from consumers. It can be said that the consumers' perceived trust lacks a foundation, as pointed out by many studies that without a reliable outside monitoring system, all the signals provided by suppliers could just be "cheap talk" (McCluskey, 2000; Janssen and Hamm, 2012; for exam-

ple). As such, without a credible government action, the trust consumers put on the signals will eventually fade, and the market for safe food can not be sustained. Hence, building up the trust in governmental management is crucial.

## APPENDIX

### Appendix 1: Test for non-parametric model

Test Statistic 'Jn': 0.1380852      P Value: 0.077694 .  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
Null of correct specification is rejected at the 10% level

### Appendix 2: Test for statistical significance of variables

Individual Significance Tests  
P Value:  
trust < 2.22e-16 \*\*\*  
type < 2.22e-16 \*\*\*  
consumption < 2.22e-16 \*\*\*  
educ < 2.22e-16 \*\*\*  
google 0.0050125 \*\*  
children < 2.22e-16 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

### Acknowledgement:

This work was financially supported by National Foundation for Science and Technology (NAFOSTED) Vietnam through project 502.01-2017.13. We would like to express our thanks to the financial support. We also thanks to anonymous referees for their helpful comments.

### Notes:

1. <http://mobitv.net.vn/tin-avg/201605/Thi-truong-rau-an-toan-Khi-cung-cau-khong-gap-nhau-14218/>
2. MARD (2008), Good agricultural practices for production of fresh fruit and vegetables in Vietnam (VietGAP)
3. <http://www.thesaigontimes.vn/138886/Sau-7-nam-dien-tich-trong-rau-VietGap-moi-dat-04.html>

---

## References

- Akerlof, G. (1970), 'The market for 'lemons': Qualitative uncertainty and the market mechanism', *Quarterly Journal of Economics*, 84, 488-500.
- Alexander, H. Wertheim (2014), *Perception & control of self – motion*, Institute for Perception TNO Soesterberg, The Netherlands.
- Chen, M.F. (2007), 'Consumer attitudes and purchase intentions in relation to organic foods in Taiwan: Moderating effects of food-related personality traits', *Food Quality and Preference*, 18(7), 1008-1021.
- Chih-Ching Teng and Yu-Mei Wang (2015), 'Decisional factors driving organic food consumption: Generation of consumer purchase intentions', *British Food Journal*, 117(3), 1066-1081.
- Gracia, A. and Magistris, T.D. (2008), 'The demand for organic foods in the south of Italy: a discrete choice model', *Food Policy*, 33(5), 386-396.
- Hsiao E. Y, Conley.A, Howell. D.A, Sullivan.A, Pritchett. C.J, Carlberg. R.G, Nugent. P.E, Phillips. M.M (2007), 'K-Corrections and Spectral Templates of Type Ia Supernovae', *The Astrophysical Journal*, 663 (2), 1187-1200.
- Janssen, M. and Hamm, U. (2012), 'Product labelling in the market for organic food: consumer preferences and willingness-to-pay for different organic certification logos', *Food Quality and Preference*, 25(1), 9-22.
- Li, Q. and Racine, J.S. (2004), 'Cross- validated local linear nonparametric regression', *Statistica Sinica*, 14, 485-512.
- McCluskey, J.J. (2000), 'A Game Theoretic Approach to Organic Foods: An Analysis of Asymmetric Information and Policy', *Agricultural and Resource Economics Review*, 29(1), 1-9.
- Mergenthaler, M., Weinberger, K. and Qaim, M. (2009), 'The role of consumers' perceptions in the valuation of food safety and convenience attributes of vegetables in Vietnam', presentation at the *International Association of Agricultural Economists Conference*, IAAE, August 16<sup>th</sup>-22<sup>th</sup>.
- Minh, N.T, Anh. L.T, Hue. T.N (2017), 'Vietnamese consumers' behavior and willingness to pay for safe vegetable in the presence of information asymmetry – study for Hanoi', *Vietnam Journal of Science, Technology and Engineering*, 19(8), 43-48.
- Nadaraya, E.A. (1964), 'On Estimating Regression', *Theory of Probability and its Applications*, 9(1), 141-200, DOI:10.1137/1109020.
- Nguyen Thi Hong Trang (2016), '*Factors affecting application of Good Agricultural Practices for vegetable farming in Vietnam*', National Economics University.
- Nguyen Thi Minh and Hoang Bich Phuong (2012), 'The impact of asymmetric information in Vietnam's health insurance', *Journal of Economics and Development*, 14 (3), 5-21.
- Nguyen Thi Minh, Trinh Trong Anh, Nguyen Hong Nhat, Vu Thi Bich Ngoc and Lam Van Son (2014), 'Impacts of state ownership on asymmetric information at listed firms: a case study of HOSE', *Journal of Economics and Development*, 202, 60-67.
- Roosen J., Lusk, J.L. and Fox, J.A. (2003), 'Consumer Demand for and Attitudes Toward Alternative Beef Labeling Strategies in France, Germany, and the UK', *Agribusiness*, 19(1), 77-90.
- Spence, M. (1973), 'Job Market Signaling', *Quarterly Journal of Economics*, 87(3), 355-374.
- Stiglitz, Joseph E. (1975), 'The Theory of "Screening", Education and Distribution of Incomes', *The American Economic Review*, 65 (3), 283-300.
- Watson, G.S. (1964), 'Smooth regression analysis', *The Indian Journal of Statistics*, 26(4), 359-372.
- Xu, L and Wu, L. (2010), 'Food safety and consumer willingness to pay for certified traceable food in China. J', *Sci. Food Agric*, 90, 1368-1373.